Do electric researchers dream of strategy?

The limits and desirability of automated data analysis

Jason Leman
Surveys Officer
Higher Education Academy

Session summary

> Can Institutional Research be automated?

> Can all kinds of data analysis be automated?

> A practical example

> Implications for Institutional Research

The automation of analysis

The electric researcher is born

The automation of analysis

> Probability that **statisticians** will be automated: **22%**

"Develop or apply mathematical or statistical theory and methods to collect, organize, interpret, and summarize numerical data to provide usable information."

> Probability that **survey researchers** will be automated: **23%**

"Plan, develop, or conduct surveys. May analyze and interpret the meaning of survey data, determine survey objectives, or suggest or test question wording"

> Probability that **mathematical technicians** will be automated: **99%**

"Apply standardized mathematical formulas, principles, and amethodology to technological problems in engineering and physical sciences in relation to specific industrial and research objectives, processes, equipment, and products"
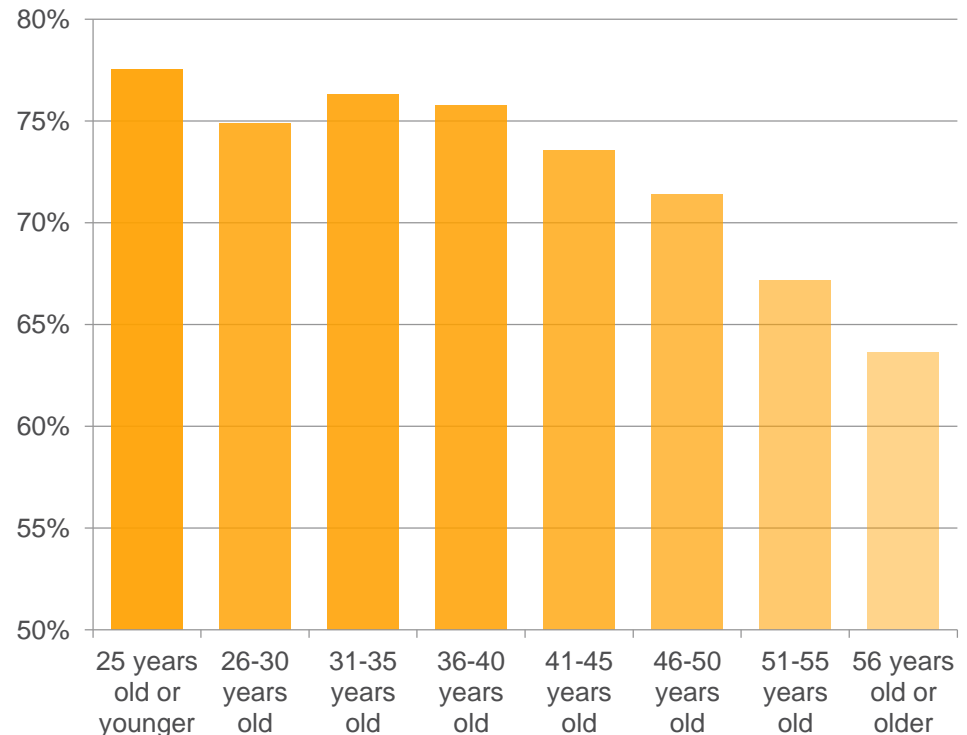
(Frey and Osbourne 2013)

The automation of analysis

> ## Descriptive analysis…

- Descriptive analysis is often the first step of analysis, setting out broad relationships in the data

- Automation of analysis often starts with descriptive analysis

- Without context, the meaning of these relationships may be mis-understood

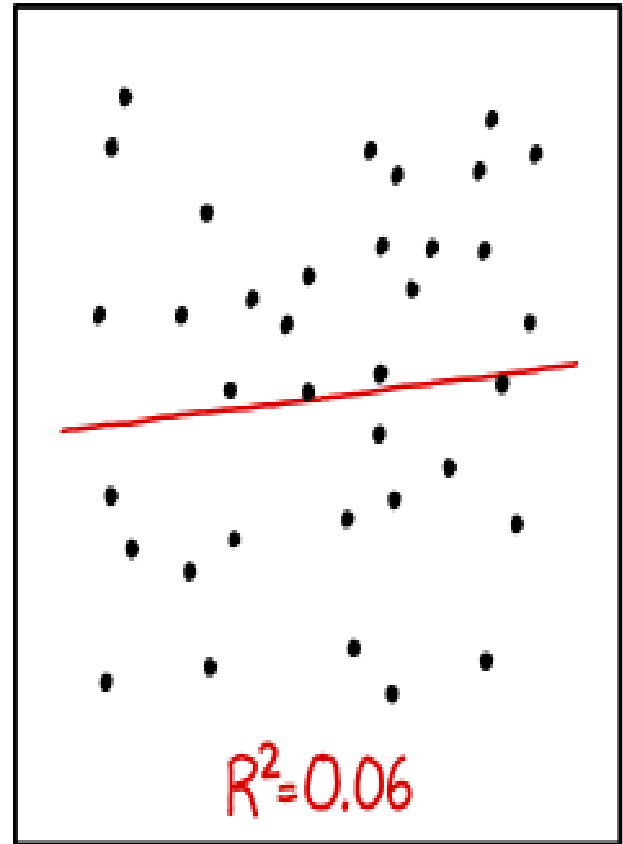- Without statistical testing, spurious differences may be given credance

Encouraged to think about devoloping career skills

The automation of analysis
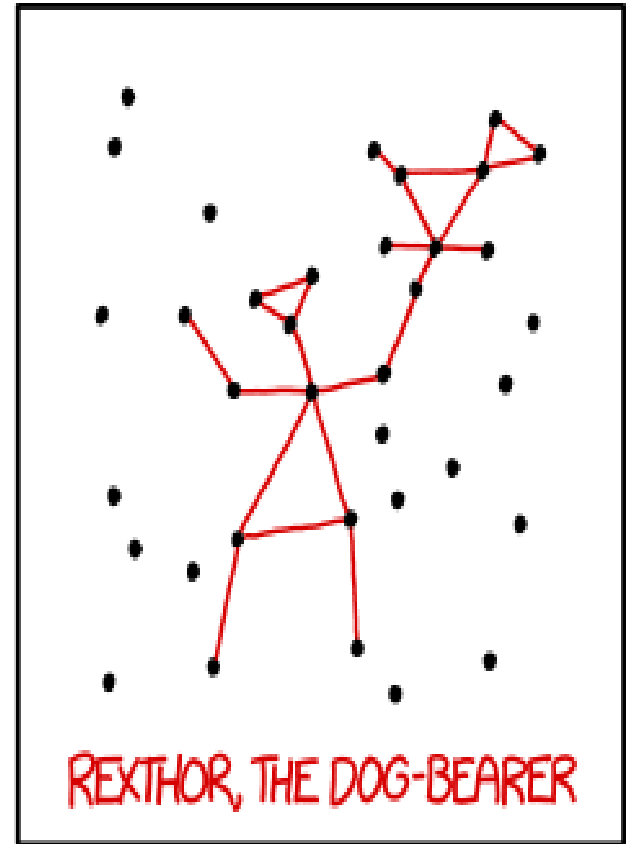
> **Regression and modelling**

- Regression analysis may be used to evaluate the impact of factors upon particular measures

- Survey research has an issue in that factors are often closely correlated, which makes regression analysis less reliable in estimating true relationships

- Machine modelling of data is sensitive to how variables are described and used - garbage in, garbage out

$R^2 = 0.06$

The automation of analysis

> **Regression and modelling**

- Regression analysis may be used to evaluate the impact of factors upon particular measures

- Survey research has an issue in that factors are often closely correlated, which makes regression analysis less reliable in estimating true relationships

- Machine modelling of data is sensitive to how variables are described and used - garbage in, garbage out

REXTHOR, THE DOG-BEARER
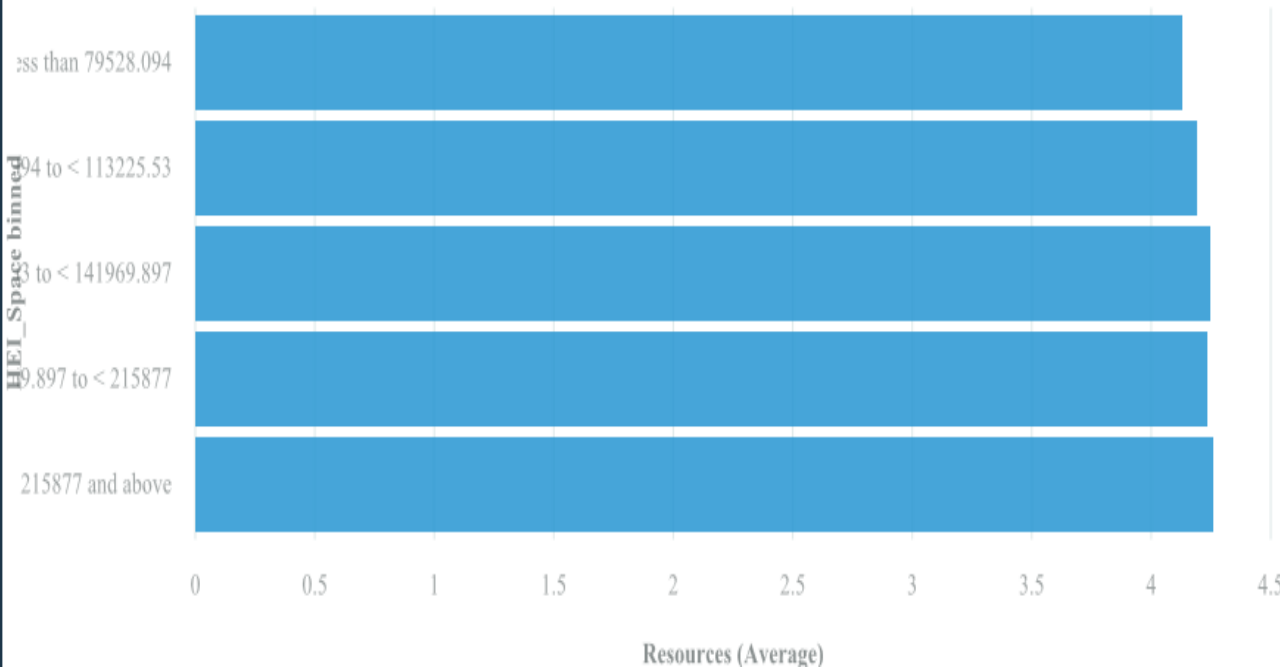
(XKCD)

Watson Analytics enables sophisticated analysis of data 'at the push of a button', but does not inform the user as to the research process that underlies that analysis. The user is presented with an 'answer', rather than with a nuanced picture of uncertainty.
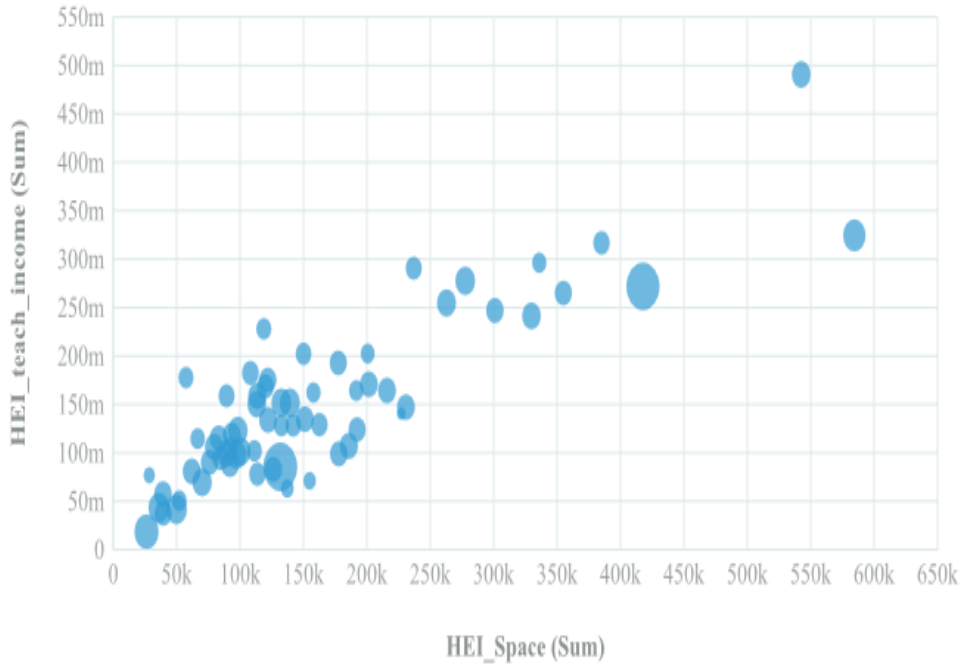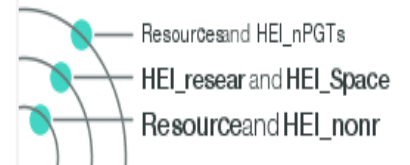
The tool is useful for the easy exploration and visualisation of data, but the user needs awareness of what such relationships mean to draw robust conclusions

eyondCore also offer analysis 'at the touch of a button', but in the smallprint make clear that thout a good understanding of the data, results are likely to be meaningless. Whilst these ckages are promoted as needing little expertise, there is a risk of 'garbage in – garbage ou

Rows represent Transactions   and the unit is Units

(Purchases, Patient Visits, Transactions, ...)   (Dollars, Days, Returns, ...)

12 **What do the rows represent?**

BeyondCore only knows what is in your data, so it doesn't know what each row represents. Each row might be a customer, a patient visit, a purchase, a shipment, etc. Here you specify what BeyondCore should call each row in the story.

⚠ When you upload your own data to BeyondCore, be very careful about the structure of your data. Each row should be at the right granularity for your analysis. For example, if I want to analyze Customer Lifetime Value, each row should be a customer, and the outcome would be the total amount of purchases by the customer. But if I wanted to understand the pattern of purchases in general, each row would be a purchase and the variable I want to understand would be the purchased amount.

Next

BeyondCORE

WINLOSS

WON BY REGION WHEN COMPETITOR IS ORACLE CORPORATION

■ only Oracle Corporation   ▨ all other COMPETITOR

South East Asia when Oracle Corporation
Difference From Overall: 15.8
Difference From Usual: 24.87 (162.1%)
Average: 40.22
Standard Deviation: 49.03
Count: 276
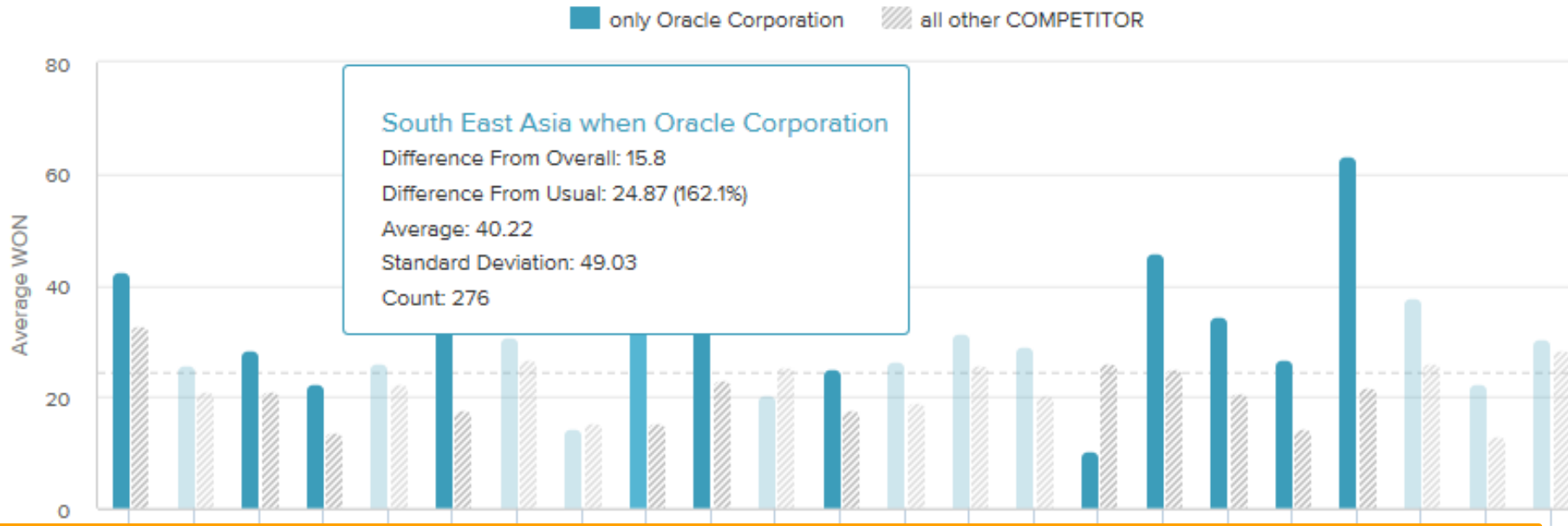
As with Watson, BeyondCore offers good options around the visualisation of data. There is also some contextual information on statistical tests and pointers towards analytical strategies.

When COMPETITOR is Oracle Corporation, REGION: Midwest and Channel do better

But you should also be aware of the case when COMPETITOR is Oracle Corporation, because the WON demonstrated a different pattern in this case. If you compare the two graphs, you

The automation of analysis

> Factor analysis and SEM…

Factor analysis is used to simplify analysis and dissemination through exploring higher level factors. Multiple items may be combined into scales and scalelets that are typically more reliable as measures than individual items.

Factors are often built into questionnaire design, in which case Structural Equation Modelling is more appropriate. However, SEM is more complex, and thus is typically underused.

"Exploratory factor analysis (EFA) is rightly described as both an art and a science, where researchers follow a series of analytic steps involving judgments more reminiscent of qualitative inquiry, an interesting irony given the mathematical sophistication underlying EFA models."

Osborne et al (2008)

The automation of analysis

> Factor analysis…

	Component			Component			Component			
	1	2		1	2		1	2	3	4
Q16.a Writing…	0.179	0.631		0.72	0.013		-0.059	-0.013	-0.011	-0.919
Q16.b Speaking…	0.336	0.507		0.621	-0.129		0.235	-0.142	0.004	-0.449
Q16.c Thinking…	0.12	0.697		0.778	0.065		0.26	-0.016	0.207	-0.365
Q16.d Analysing…	-0.226	0.789		0.511	0.065		-0.013	-0.022	0.682	-0.021
Q16.e Employability…	0.053	0.645		0.531	-0.054		0.258	-0.082	0.344	-0.049
Q16.f Independent…	0.189	0.561		0.6	-0.045		0.507	0.009	0.119	-0.116
Q16.g Innovative…	0.405	0.388		0.499	-0.201		0.767	-0.012	-0.047	-0.016
Q16.h Working…	0.459	0.314		0.402	-0.283		0.514	-0.183	0.051	0.01
Q16.i Developing…	0.826	0.032		0.087	-0.735		0.207	-0.657	-0.065	-0.019
Q16.j Understanding…	0.912	-0.119		-0.1	-0.876		0.003	-0.825	-0.11	-0.03
Q16.k Exploring…	0.74	0.084		0.121	-0.654		-0.043	-0.724	0.167	0.003
Q16.l Active…	0.807	0.051		0.071	-0.752		-0.026	-0.783	0.069	-0.036
	PCA			PAF			PAF - extracted four factors			

Types of analysis like factor analysis require not just statistical techniques, but a research approach that incorporates judgements around what is meaningful and desirable. Automated analysis may make such processes easier, but need to enable informed research choices rather than presenting results as concrete.

4   Model criticism

In this section I have attempted to falsify the model that I have presented above to understand what aspects of the data it is not capturing well. This has been achieved by comparing the model with data I held out from the model fitting stage. In particular, I have searched for correlations and dependencies within the data that are unexpectedly large or small. I have also compared the distribution of the residuals with that assumed by the model (a normal distribution). There are other tests I could perform but I will hopefully notice any particularly obvious failings of the model. Below are a list of the discrepancies that I have found with the most surprising first. Note however that some discrepancies may be due to chance; on average 10% of the listed discrepancies will be due to chance.

Systems such as "The Automatic Statistician", now funded by Google, raise the possibility of advanced analysis, standardised research strategies, and text outputs. Here, the programme has created a paper that describes research choices and statistical uncertainty in plain English.

a)   b)

Lloyd (2014) The Automatic Statistician
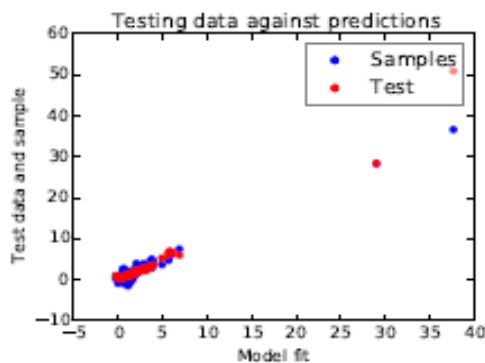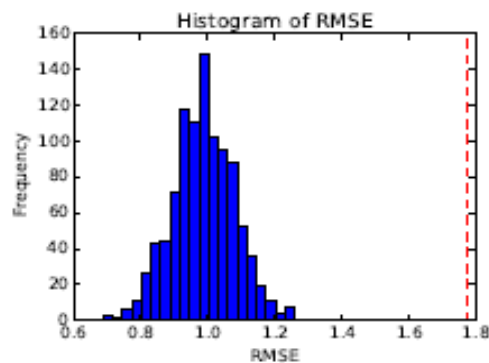
The automation of analysis

> There is increasing competition for "easy to use" analytical software, with statistics program moving to cater for easier analysis

> Programs like IBM Watson analytics offer analysis "at the push of a button", guiding automated data analysis through machine learning and statistical analysis, including factor analysis, neural networks, and regression

> Proposals for student surveys that automate factor analysis to produce "on the fly" scales

> There is increasing potential for smart data analytics, offering sophisticated data analysis alongside embedding multiple kinds of evidence alongside quantitative measures

> Analysis "at the push of a button" risks poor decisions being made

> Good smart analytics programmes guide users through research choices and inform them of the fuzziness of findings

An automated reporting engine

Making my own electric researcher
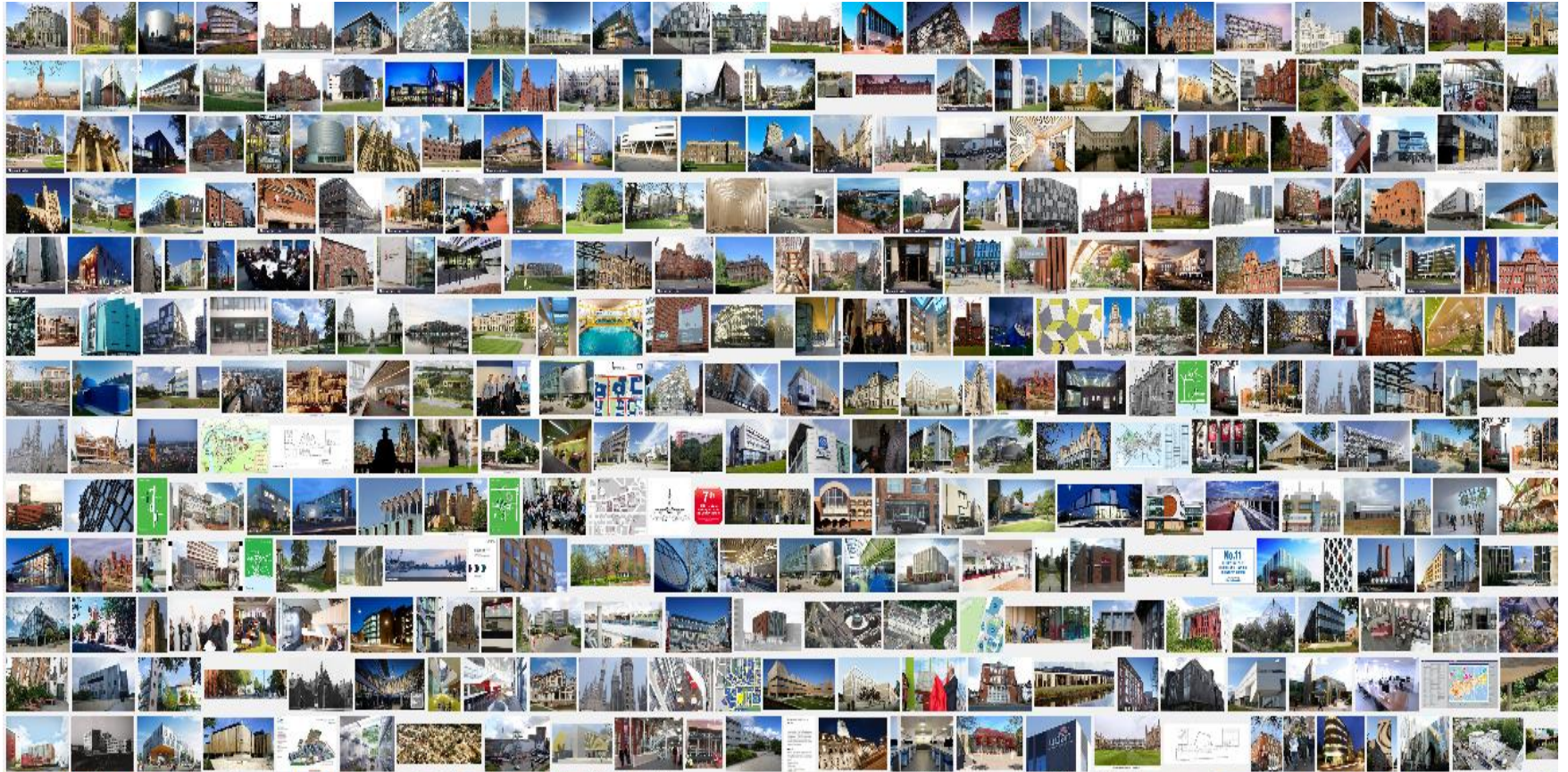
An automated reporting engine

My own development of an automated reporting engine began with the need to handle substantial amounts of information for a single large institution. There was a lot of data to get through!

An automated reporting engine

Initial reporting needs were driven by the need for multiple statistical tests across numerous breakdowns to identify where important differences exist.

Chi-square was used as a relatively robust test to identify important differences where response numbers varied widely.

However, significance levels of $p<0.05$ often produced too many flags, whilst $p<0.01$ too few. Development needed to incorporate more sophistication.

An automated reporting engine

I then moved from a single institution to the Higher Education Academy, where reports were needed for over 150 institutions. This created an additional challenge.

An automated reporting engine
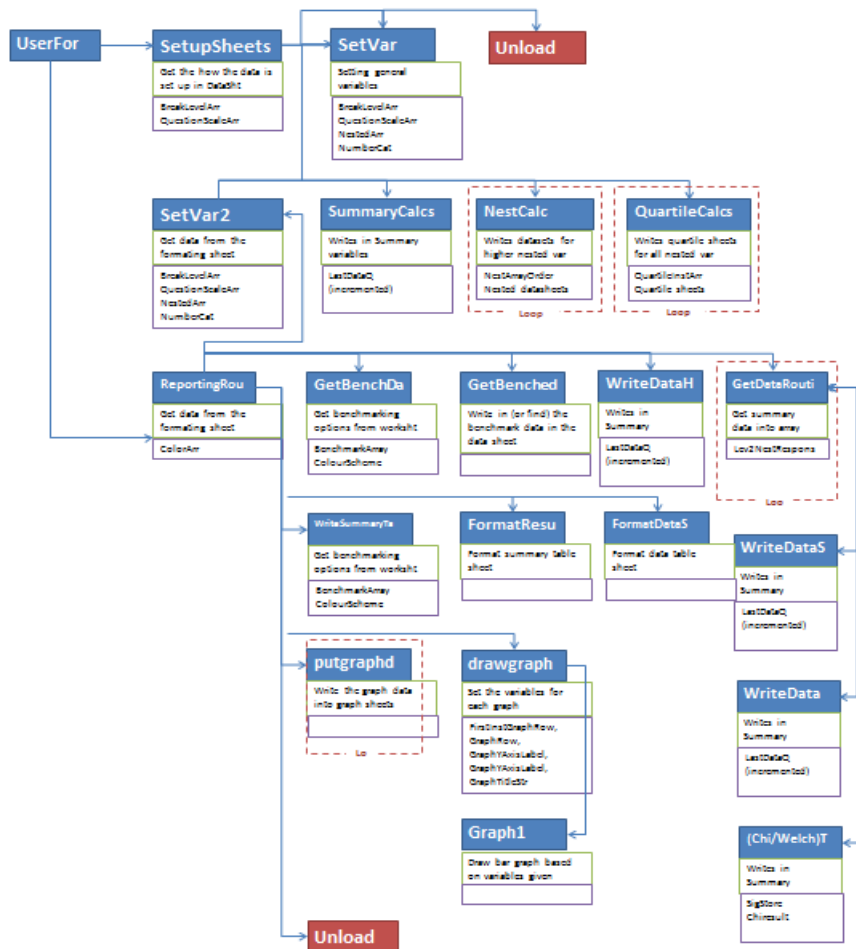
The engine has been rewritten twice, with increasing flexibility, capability, and integration.

The engine is written in VBA, separate stages of reporting processes being handled by largely separate routines.

Currently the engine has 3747 lines of code, with six essentially separate modules covering the process from raw data to report production.

An automated reporting engine

Quartile and weighted calculations are pre-calculated prior to reporting. Sheets can be filtered to produce quartile or effect size heat maps.

Effect sizes were used to define the colouring of significant cells. This meant that many significance flags were created, but prioritised through colour coding of cells according to effect size. The Holm–Bonferroni adjustment for multiple chi-squared tests was added to reduce false positives.

The expansion of statistical calculations involved in the creation of each report, along with the need for overviews of statistical calculations, prompted a move away from large arrays. Multiple temporary sheets were used to store interim calculations instead.

	A	C	D	E	F	G	H	I	J	K	L	M	N
1						Summary							
2	PTES 2016 results	Average Response	Teaching	Engagement	Assessment	Dissertation	Organisation	Resources	Skills development	Overall	Information		Q2_1_a Staff are good at explaining things
13	**All responses**												
14	**Gender**												
15	Male	31684	82%	79%	72%	80%	75%	84%	79%	83%	84%		88%
16	Female	47384	83%	79%	74%	78%	74%	84%	77%	82%	85%		**88%**
17	**Disability**												
18	Disabled	6149	79%	75%	71%	75%	69%	82%	73%	78%			
19	No known disability	71140	83%	79%	74%	79%	74%	85%	78%	83%			
20	**Fee status**												
21	UK	38080	83%	79%	75%	79%	72%	84%	77%	83%			
22	Other European Union (EU)	10072	78%	75%	66%	75%	70%	83%	73%	78%			
23	Non-EU	31321	82%	79%	74%	80%	77%	85%	80%	83%			
24	**Nationality**												
25	UK incl. Channel Islands	41588	83%	79%	74%	79%	73%	84%	77%	83%	85%		88%
26	Asia	15410	84%	80%	77%	81%	79%	86%	82%	85%	86%		90%
27	Other Europe	1125	79%	75%	67%	73%	73%	84%	74%	79%	84%		86%
28	Africa	3568	87%	82%	77%	82%	81%	86%	87%	89%	90%		91%

78% (n=6363) is significantly lower than 83% (n=82476). Chi-Square (2,88839) = 99.81, p = 0.000, phi = 0.33

Summary for Poppleton University compared to the Sector benchmark.

Poppleton University to Sector quartiles

Postgraduate Taught Experience Survey

Higher Education Academy

79%
75%
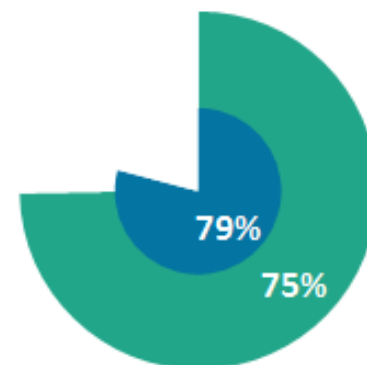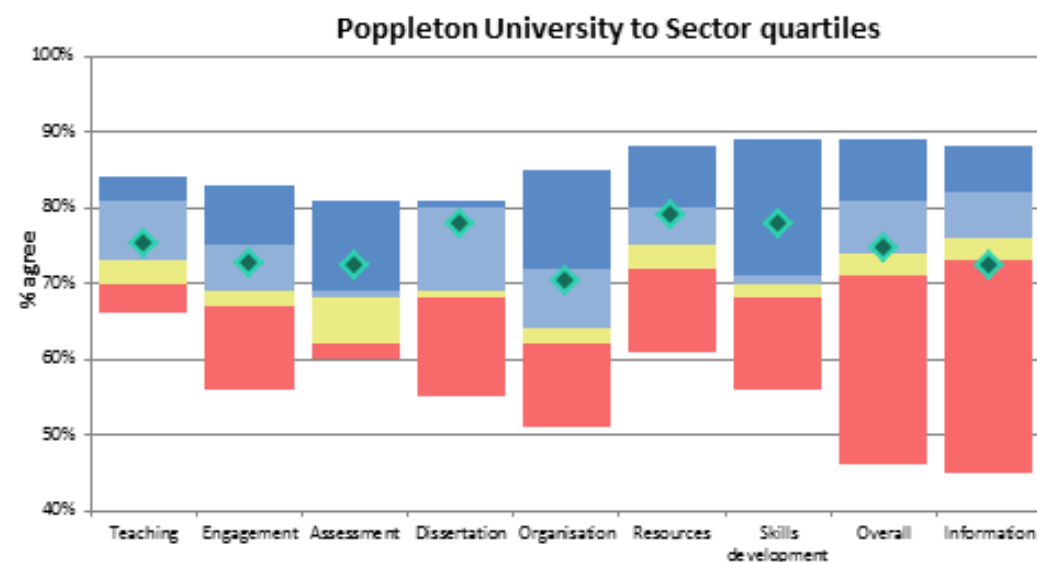
● **Poppleton University**
● **Sector**

Aggregate experience score

Key: Top quartile   Lower quartile   ◆ Poppleton University
Upper quartile   Bottom quartile

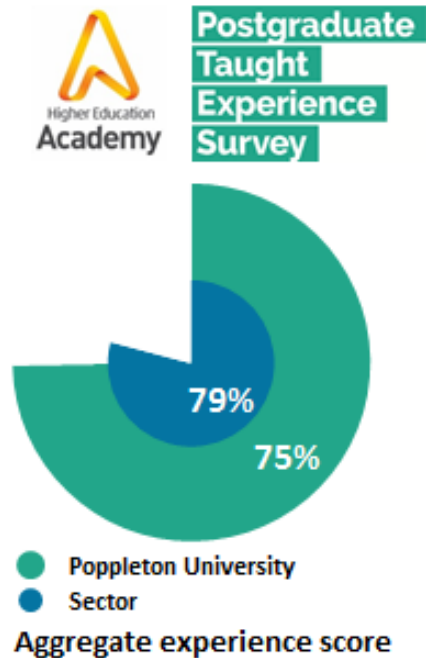|	Poppleton University	Sector	Difference	Significance
Comparing Poppleton University to the Sector benchmark, the most positive statements were:				
1: There is sufficient contact time to support effective learning (n = 157)	79%	69%	10.2%	p<0.05
2: As a result of the course I am more confident about independent learning (n = 158)	87%	82%	4.8%	
3: My ability to communicate information effectively to diverse audiences has developed during my course (n = 157)	80%	75%	4.5%	
Comparing Poppleton University to the Sector benchmark, the most negative statements were:				
1: The workload on my course has been manageable (n = 157)	54%	72%	-18.0%	p<0.001
2: The course is intellectually stimulating (n = 156)	69%	85%	-16.8%	p<0.001
3: Information for prospective students was accurate (n = 152)	65%	80%	-15.1%	p<0.001

The quartile graph shows the range of institutional scores in PTES for each measure, divided into quartiles. For example, the lower quartile covers the range from the lowest scoring institution across a quarter of the lower scoring institutions. The diamond indicates where this institution is on the scale. The overall satisfaction graph shows the institution and benchmark scores for the overall question. The positive and negative statements are the three statements (where an agree/diagree scale is used) with the largest positive and negative differences.

An automated reporting engine

Postgraduate Taught Experience Survey

79%
75%

● Poppleton University
● Sector
Aggregate experience score

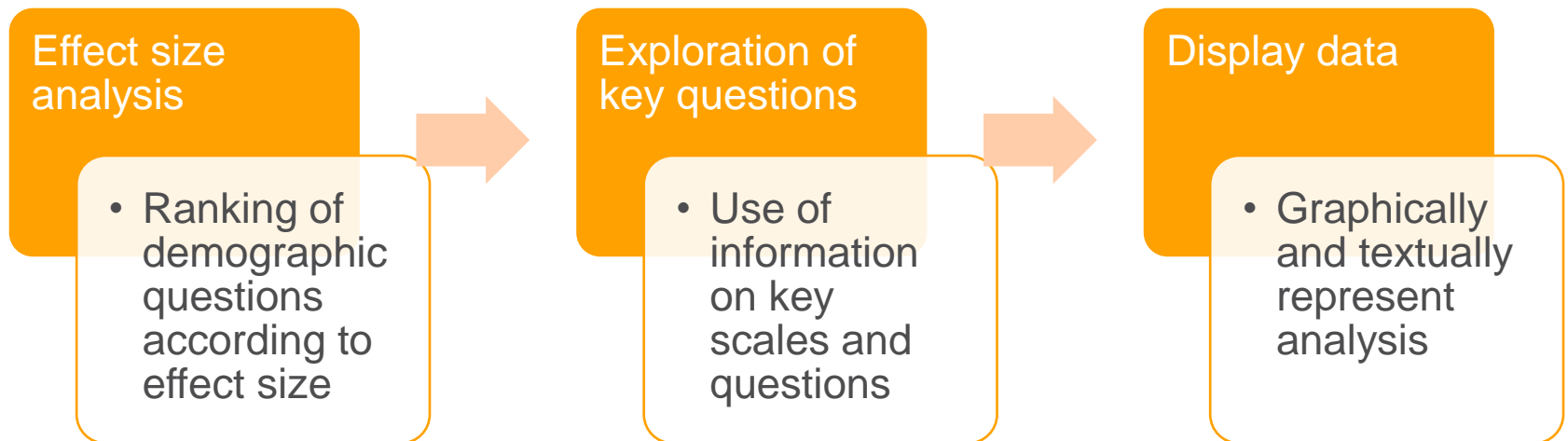From summary tables the next step was a dashboard that sought to enable clear prioritisation of action.

This process begun with graphical representations of relative position.

It was then developed through textual descriptions of the largest differences and the listing of priority questions related to those differences.

However, when it came to displaying demographic analyses, a method was needed to narrow down the many possible questions.

An automated reporting engine

> Demographic analysis…

Effect size analysis
- Ranking of demographic questions according to effect size

Exploration of key questions
- Use of information on key scales and questions

Display data
- Graphically and textually represent analysis

One challenge was to select particular demographic questions to display. This required a prioritisation. Using a research strategy of selecting questions of interest by effect size, the engine progresses through logical steps to present key analyses.

An automated reporting engine

Difference between Poppleton and Sector upper quartile, by Ethnicity
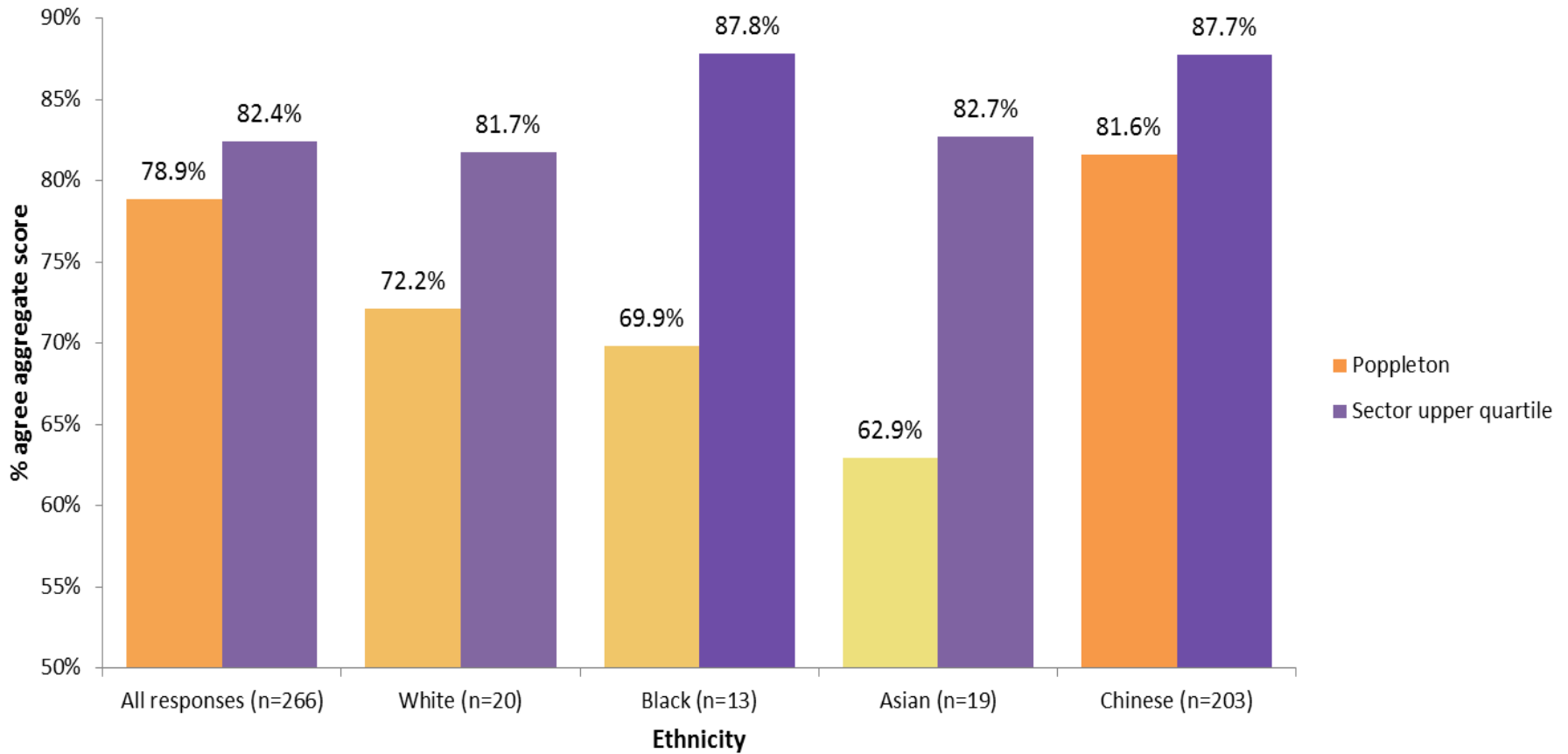
An automated reporting engine

There were relatively large differences within Ethnicity for poppleton compared to the Sector upper quartile benchmarking group. 'Chinese' had the least negative results at poppleton relative to the benchmark with an aggregate score -6.1% lower than the Sector upper quartile benchmark. This was a significant ($p<0.05$) difference. The category most negative relative to the Sector upper quartile benchmarking group was 'Asian', with a score 19.8% lower than the Sector upper quartile benchmark.

Within Ethnicity - Chinese, the most positive area of feedback was Dissertation. Here, the most positive statements were:	poppleton	Sector upper quartile	Difference	Significance
1: I understand the required standards for the dissertation / major project	93%	89%	3.5%	
2: I am happy with the support I received for planning my dissertation / major project	86%	86%	0.1%	
Within Ethnicity - Asian, the most negative area of feedback was Resources. Here, the most negative statements were:				
1: The library resources and services are good enough for my needs	37%	87%	**-49.7%**	p<0.001
2: I have been able to access subject specific resources necessary for my studies	53%	86%	-33.9%	p<0.001

Graphical, numerical and textual presentation of data increases the potential that the end user will understand and disseminate findings, regardless of their understanding of statistical methods.

An automated reporting engine

Automated processes:

Customisation of reports around benchmarking groups, colour scheme, etc

Tabular statistically tested output

Analysis based upon difference, significance and effect size

Graphical and textual representation

Non-automated process:

Aggregation of variables into grouped variables and scales

Design and format of analysis

Further alteration and adaptation

Links to enhancement and strategy – the issues and context of the institution

Objective and comprehensive use of statistics to define priorities according to data

Subjective and partial use of statistics to define priorities according to strategy

The engine does not automate many functions, such as factor analysis. End users often wish to select particular parts of analyses or otherwise adapt to fit into institutional context and enable enhancement. The engine is being developed to improve how it informs users of research choices.
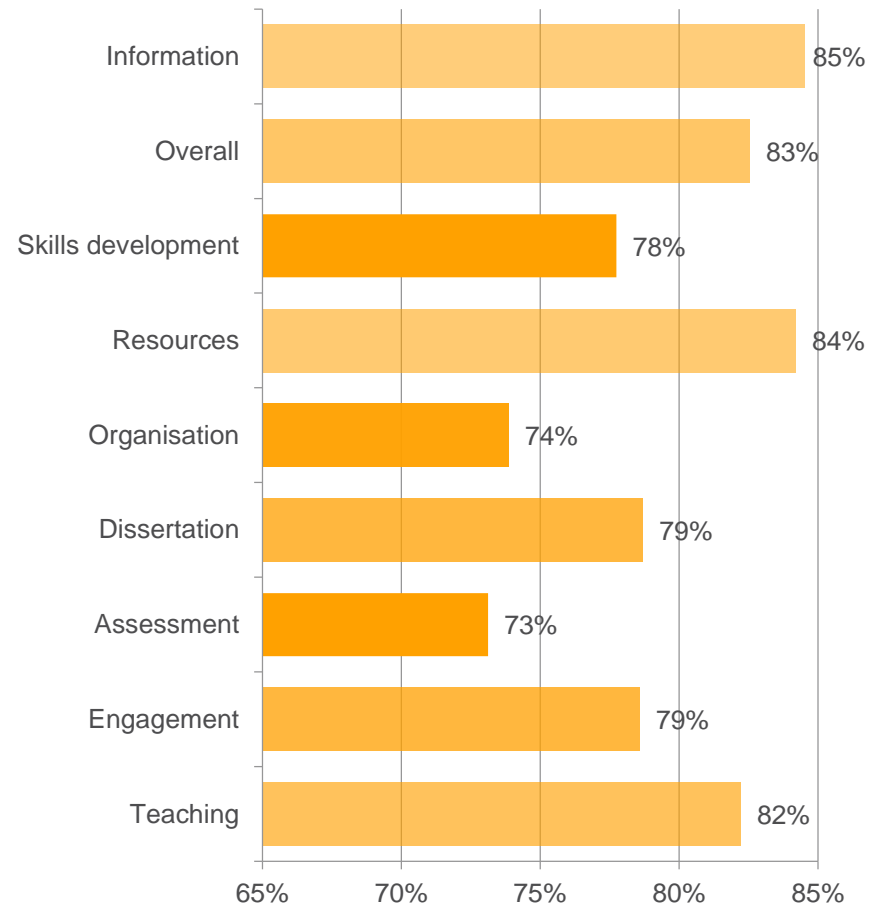
Do electric researchers dream of strategy?
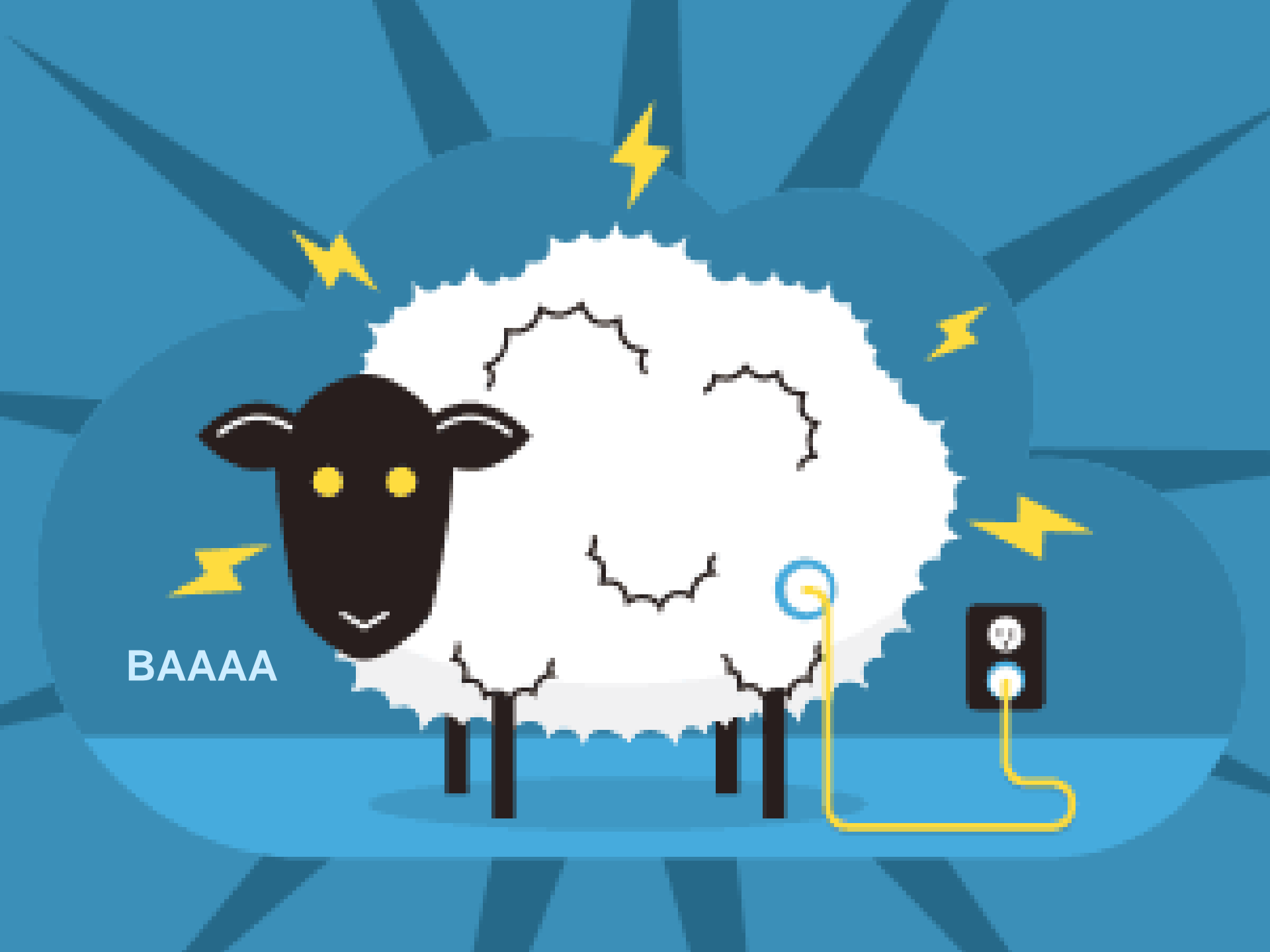
BAAAA

Do electric researchers dream of strategy?

What analysis is reported is informed by issues and context.

| Category	%
Information	85%
Overall	83%
Skills development	78%
Resources	84%
Organisation	74%
Dissertation	79%
Assessment	73%
Engagement	79%
Teaching	82%

What is the priority – the lowest scoring scales such as Assessment and Organisation?...

Do electric researchers dream of strategy?

What analysis is reported is informed by issues and context.

How does this decision impact upon dissemination, enhancement, politics?
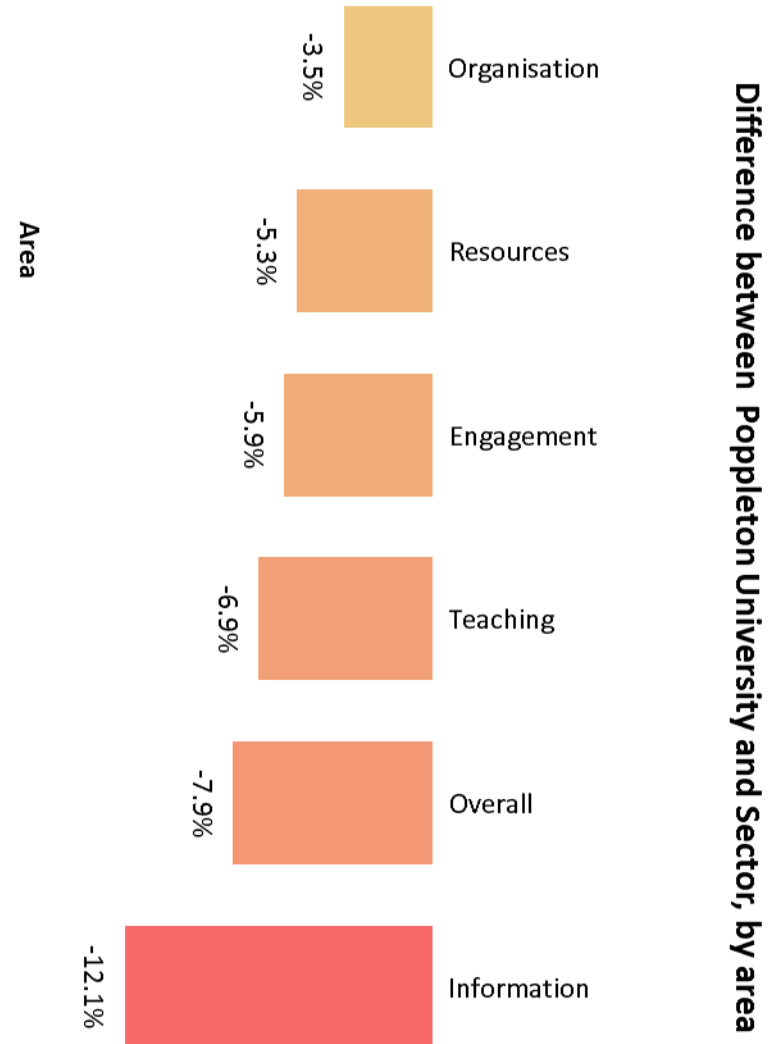
…Or is the priority the scale that scores lowest relative to other institutions?...

Do electric researchers dream of strategy?

What analysis is reported is informed by issues and context.

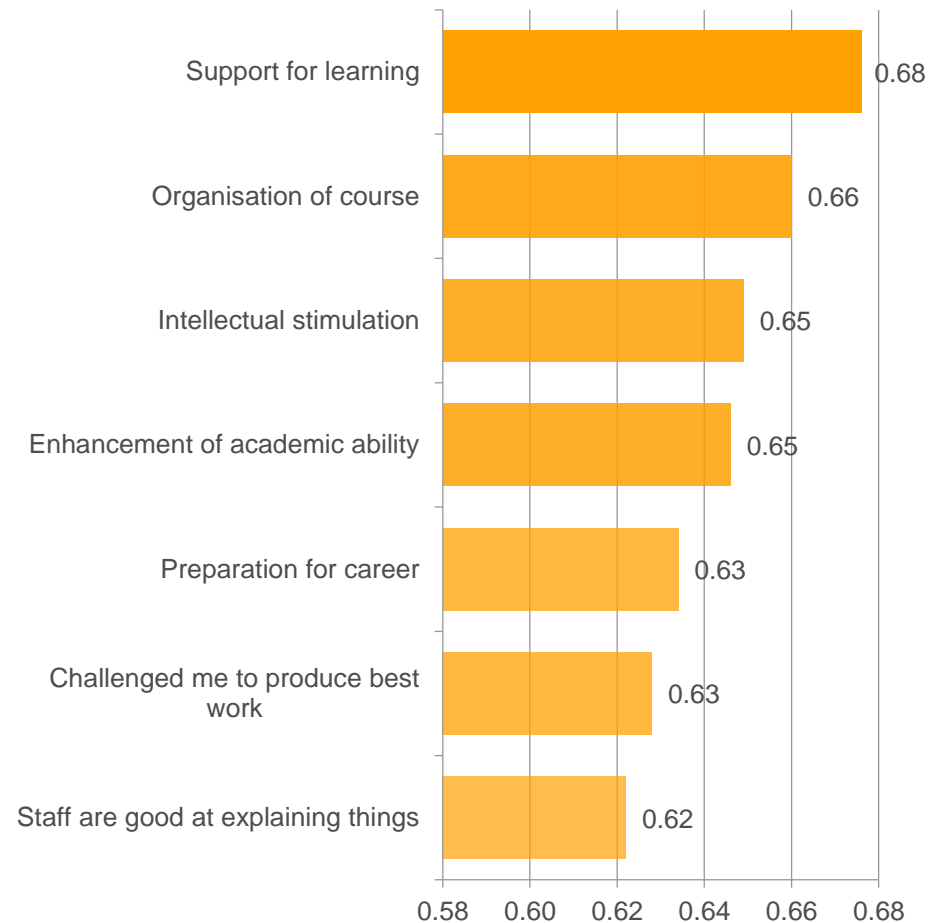How does this decision impact upon dissemination, enhancement, politics?

IR is moving up Terenzini's hierarchy, from technical analysis to issues and context.

Support for learning	0.68
Organisation of course	0.66
Intellectual stimulation	0.65
Enhancement of academic ability	0.65
Preparation for career	0.63
Challenged me to produce best work	0.63
Staff are good at explaining things	0.62

(x-axis: 0.58 0.60 0.62 0.64 0.66 0.68)

…Or is the priority the area that is most correlated with the overall experience? Whilst automation can make this analysis easier, the choice cannot be automated.

Do electric researchers dream of strategy?

What analysis is reported is informed by issues and context.

How does this decision impact upon dissemination, enhancement, politics?

IR is moving up Terenzini's hierarchy, from technical analysis to issues and context.

IR connects to theory, strategy, and teaching & learning practice – across departments, institutions, and the sector.

Fulfilling our Potential: Teaching Excellence, Social Mobility and Student Choice

Presented to Parliament by the Secretary of State for Business, Innovation and Skills by Command of Her Majesty

November 2015

The value of automated analysis is in giving people more sophisticated tools to find differences that matter…
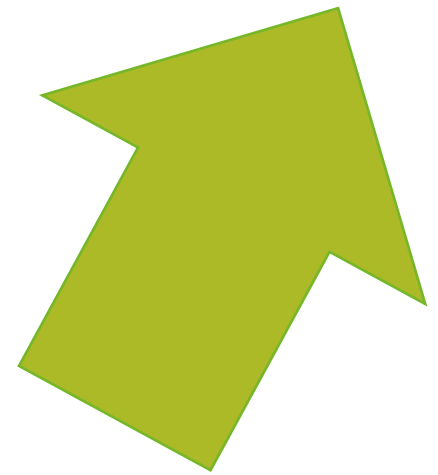
…the value of the institutional researcher is increasingly about bridging this analysis to the context and issues institutions face.

References

The Automatic Statistician http://www.automaticstatistician.com/

Frey, C. B., & Osborne, M. A. (2013). The future of employment: how susceptible are jobs to computerisation. http://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

Hammond, N. (2009) Institutional research and the student learning experience, Perspectives: Policy and Practice in Higher Education, 13:3, 80-83

King, B. R. (2016). Making the Connections Across Institutional Types and Academic Programs: Recommendations for Institutional Research Practice and Future Research. *New Directions for Institutional Research*, *2015*(168), 101-105.

Volkwein, J. F. (1999). The four faces of institutional research. *New directions for institutional research*, *1999*(104), 9-19.

Wells, R. S., & Stage, F. K. (2015). Past, Present, and Future of Critical Quantitative Research in Higher Education. *New Directions for Institutional Research*, *2014*(163), 103-112.

Questions?

jason.leman@heacademy.ac.uk   @HEA_surveys